

THE RADICAL OF THE KERNEL OF A CERTAIN DIFFERENTIAL OPERATOR AND APPLICATIONS TO LOCALLY ALGEBRAIC DERIVATIONS

WENHUA ZHAO

ABSTRACT. Let R be a commutative ring, \mathcal{A} an R -algebra (not necessary commutative) and V an R -subspace or R -submodule of \mathcal{A} . By the *radical* of V we mean the set of all elements $a \in \mathcal{A}$ such that $a^m \in V$ for all $m \gg 0$. We derive (and show) some necessary conditions satisfied by the elements in the radicals of the kernels of some (partial) differential operators such as all differential operators of commutative algebras; the differential operators $P(D)$ of (noncommutative) \mathcal{A} with certain conditions, where $P(\cdot)$ is a polynomial in n commutative free variables and $D = (D_1, D_2, \dots, D_n)$ are either n commuting locally finite R -derivations or n commuting R -derivations of \mathcal{A} such that for each $1 \leq i \leq n$, \mathcal{A} can be decomposed as a direct sum of the generalized eigen-subspaces of D_i ; etc. We then apply the results mentioned above to study R -derivations of \mathcal{A} that are locally algebraic or locally integrable over R . In particular, we show that if R is an integral domain of characteristic zero and \mathcal{A} is reduced and torsion-free as an R -module, then \mathcal{A} has no nonzero locally algebraic R -derivations. We also show a formula for the determinant of a differential vandemonde matrix over commutative algebras. This formula not only provides some information for the radicals of the kernels of ordinary differential operators of commutative algebras, but also is interesting on its own right.

1. Background and Motivation

Let R be a commutative ring and \mathcal{A} an R -algebra (not necessary commutative). A *derivation* D of \mathcal{A} is a map from \mathcal{A} to \mathcal{A} such that $D(ab) = D(a)b + aD(b)$ for all $a, b \in \mathcal{A}$. If D is also R -linear, we call it an *R -derivation* of \mathcal{A} .

Date: January 24, 2017.

2000 Mathematics Subject Classification. 47F05, 47E05, 16W25, 16D99.

Key words and phrases. The radical, the kernel of a differential operator, locally algebraic or integral derivations, a differential vandemonde determinant, Mathieu subspaces (Mathieu-Zhao spaces).

The author has been partially supported by the Simons Foundation grant 278638.

For each $a \in \mathcal{A}$, we denote by ℓ_a the map from \mathcal{A} to \mathcal{A} that maps $b \in \mathcal{A}$ to ab . We call the associative algebra generated by ℓ_a ($a \in \mathcal{A}$) and all derivations of \mathcal{A} the *Weyl algebra* of \mathcal{A} , and denote it by $\mathcal{W}(\mathcal{A})$. The subalgebra of $\mathcal{W}(\mathcal{A})$ generated by ℓ_a ($a \in R$) and all R -derivations of \mathcal{A} will be denoted by $\mathcal{W}_R(\mathcal{A})$. Elements of $\mathcal{W}(\mathcal{A})$ are called *differential operators* of \mathcal{A} .

For each $\Phi \in \mathcal{W}(\mathcal{A})$, it is well-known and also easy to check that there exist some derivations $D = (D_1, D_2, \dots, D_n)$ of \mathcal{A} and a polynomial $P(\xi) \in \mathcal{A}[\xi]$ in n noncommutative free variables $\xi = (\xi_1, \xi_2, \dots, \xi_n)$ such that $\Phi = P(D)$, where $P(D)$ throughout this paper is defined by first writing all the coefficients of $P(\xi)$ on the left and then replacing ξ_i by D_i for all $1 \leq i \leq n$. Furthermore, if $\Phi \in \mathcal{W}_R(\mathcal{A})$, the same is true with D_i ($1 \leq i \leq n$) being R -derivations of \mathcal{A} and $P(\xi) \in R[\xi]$. We call the differential operator $\Phi = P(D)$ an *ordinary differential operator* of \mathcal{A} , if $P(\xi)$ is univariate, and a *partial differential operator* of \mathcal{A} if $P(\xi)$ is multivariate.

Next, we recall the following two notions of associative algebras that were first introduced in [Z2, Z3].

Definition 1.1. *An R -subspace V of an R -algebra \mathcal{A} is said to be a Mathieu subspace (MS) of \mathcal{A} if for all $a, b, c \in \mathcal{A}$ with $a^m \in V$ for all $m \geq 1$, we have $ba^m c \in V$ for all $m \gg 0$.*

Note that a MS is also called a *Mathieu-Zhao space* in the literature (e.g., see [DEZ, EN, EH], etc.), as suggested by A. van den Essen [E2].

The introduction of this notion is mainly motivated by the study in [M, Z1] of the well-known Jacobian conjecture (see [Ke, BCW, E1]). See also [DEZ]. But, a more interesting aspect of the notion is that it provides a natural but highly non-trivial generalization of the notion of ideals.

Definition 1.2. [Z3, p. 247] *Let V be an R -subspace (or a subset) of an R -algebra \mathcal{A} . We define the radical $\mathfrak{r}(V)$ of V to be*

$$(1.1) \quad \mathfrak{r}(V) := \{a \in \mathcal{A} \mid a^m \in V \text{ for all } m \gg 0\}.$$

When \mathcal{A} is commutative and V is an ideal of \mathcal{A} , $\mathfrak{r}(V)$ coincides with the radical of V . So this new notion is also interesting on its own right. It is also crucial for the study of MSs. For example, it is easy to see that *every R -subspace V of an R -algebra \mathcal{A} with $\mathfrak{r}(V) = \text{nil}(\mathcal{A})$ is a MS of \mathcal{A} , where $\text{nil}(\mathcal{A})$ denotes the set of all nilpotent elements of \mathcal{A} . We will frequently use this fact (implicitly) throughout this paper.*

Recent studies show that many MSs arise from the images of differential operators, especially, from the images of locally finite or locally nilpotent derivations, of certain associative algebras (e.g., see

[Z1, Z2, EWZ, EZ] and [Z4]–[Z7], etc.). For some MSs arisen from the kernels of some ordinary differential operators of univariate polynomial algebras over a field, see [EN, EH].

In this paper we study the radicals of the kernels of some ordinary or partial differential operators of \mathcal{A} and show that for certain differential operators Φ , the kernel $\text{Ker } \Phi$ is also a MS of \mathcal{A} . We also apply some results proved in this paper to study R -derivations of \mathcal{A} that are locally algebraic or locally integrable over R (see Definition 4.1). In particular, we show that if R is an integral domain of characteristic zero and \mathcal{A} is reduced and torsion-free as an R -module, then \mathcal{A} has no nonzero R -derivation that is locally algebraic over R (see Theorem 4.6). Furthermore, we also show a formula for the determinant of a differential vandemonde matrix over commutative algebras (see Proposition 5.1). This formula not only provides some information for the radicals of the kernels of ordinary differential operators of commutative algebras, but also is interesting on its own right.

Arrangement and Content: In Section 2, we assume that \mathcal{A} is commutative and derive some necessary conditions for the elements in the radical of the kernel of an arbitrary differential operators of \mathcal{A} (see Theorem 2.1 and Corollary 2.4). In particular, for every differential operator $\Phi \in \mathcal{W}(\mathcal{A})$ such that $\Phi 1_{\mathcal{A}}$ is not zero nor a zero-divisor of \mathcal{A} , the kernel $\text{Ker } \Phi$ forms a MS of \mathcal{A} .

In Section 3, we drop the commutativity assumption on \mathcal{A} but assume that $(R, +)$ is torsion-free and \mathcal{A} is reduced and torsion-free as an R -module. We first derive in Theorem 3.1 some necessary conditions satisfied by the elements in the radical of the kernel of a differential operator $P(D)$ of \mathcal{A} , where $P(\cdot)$ is a polynomial in n commutative free variables and $D = (D_1, D_2, \dots, D_n)$ are n commuting R -derivations of \mathcal{A} such that for each $1 \leq i \leq n$, \mathcal{A} can be decomposed as a direct sum of the generalized eigen-subspaces of D_i .

We then show in Proposition 3.6 that if R also is an integral domain of characteristic zero, then the conclusions in Theorem 3.1 also hold for the differential operators of \mathcal{A} which are multivariate polynomials in commuting locally finite R -derivations of \mathcal{A} . Finally, we show in Proposition 3.7 that similar conclusions as those in Proposition 3.6 (with the same assumptions on R and \mathcal{A}) also hold for all ordinary differential operators of \mathcal{A} . In particular, for all the differential operators Φ in Theorem 3.1 and Propositions 3.6 and 3.7 with $\Phi 1_{\mathcal{A}} \neq 0$, $\text{Ker } \Phi$ forms a MS of \mathcal{A} .

In Section 4, we apply some results proved in Sections 2 and 3 to study some properties of R -derivation of \mathcal{A} that are locally algebraic or

locally integral over R (see Definition 4.1). We first show in Theorem 4.3 that if \mathcal{A} is commutative and $(\mathcal{A}, +)$ is torsion-free, then every locally integral D of \mathcal{A} has its image in the nil-radical $\text{nil}(\mathcal{A})$ of \mathcal{A} . We also show in Theorem 4.6 that if R also is an integral domain of characteristic zero and \mathcal{A} is reduced and torsion-free as an R -module, then \mathcal{A} has no nonzero R -derivation that is locally algebraic over R .

In Section 5, we assume that \mathcal{A} is commutative and first show in Proposition 5.1 a formula for the determinant of a differential vandermonde matrix over \mathcal{A} . We then apply this formula in Proposition 5.4 to derive more necessary conditions satisfied by the elements in the radicals of the kernels of all ordinary differential operators of \mathcal{A} . we point out in Remark 5.3 that the formula derived in Proposition 5.1 can also be used to derive formulas for the determinants of several other families of matrices.

2. The Commutative Algebra Case

In this section, unless stated otherwise, R denotes a unital commutative ring, \mathcal{A} a commutative unital R -algebra and $\xi = (\xi_1, \xi_2, \dots, \xi_n)$ n noncommutative free variables. We denote by $\mathcal{A}[\xi]$ the polynomial algebra in ξ over \mathcal{A} , and ∂_i ($1 \leq i \leq n$) the \mathcal{A} -derivation $\partial/\partial\xi_i$ of $\mathcal{A}[\xi]$.

Once and for all, we fix in this section a nonzero $P(\xi) \in \mathcal{A}[\xi]$ and n R -derivations D_i ($1 \leq i \leq n$) of \mathcal{A} . Write $D = (D_1, D_2, \dots, D_n)$ and $P(\xi) = a_0 + \sum_{k=1}^d P_k(\xi)$ for some $a_0 \in \mathcal{A}$, $d \geq 1$ and homogeneous polynomials $P_k(\xi)$ ($1 \leq k \leq d$) of degree k in ξ .

For each $u \in \mathcal{A}$, we set $\nabla_D u := (D_1 u, D_2 u, \dots, D_n u)$, and call it the *gradient* of u with respect to D . When D is clear in the context, we will simply write $\nabla_D u$ as ∇u .

We define $P(D)$ and $P(\nabla u)$ by first writing $P(\xi)$ as a polynomial in ξ with all the coefficients on the most left (of the monomials), and then replacing ξ_i by D_i and $D_i u$, respectively, for each $1 \leq i \leq n$.

The main result of this section is the following

Theorem 2.1. *With the setting as above, let $u \in \mathcal{A}$ such that $u^m \in \text{Ker } P(D)$ for all $1 \leq m \leq d$. Then*

$$(2.1) \quad a_0 u^d = (-1)^d d! P_d(\nabla u).$$

Furthermore, if u^{d+1} also lies in $\text{Ker } P(D)$, then

$$(2.2) \quad a_0 u^{d+1} = 0.$$

To show the theorem above, we first need the following two lemmas. The first lemma is well-known and can also be easily verified by using

the mathematical induction, which is similar as the proof for the usual binomial formula.

Lemma 2.2. *Let $u \in \mathcal{A}$ and $\ell_u : \mathcal{A} \rightarrow \mathcal{A}$ that maps $a \in \mathcal{A}$ to ua . Denote by $\text{ad}_u : \mathcal{W}(\mathcal{A}) \rightarrow \mathcal{W}(\mathcal{A})$ that maps each $\Lambda \in \mathcal{W}(\mathcal{A})$ to $[u, \Lambda] := \ell_u \Lambda - \Lambda \ell_u$. Then for all $\Phi \in \mathcal{W}(\mathcal{A})$ and $k \geq 1$, we have*

$$(2.3) \quad (\text{ad}_u)^k(\Phi) = \sum_{i=0}^k (-1)^i \binom{k}{i} u^{k-i} \Phi \circ \ell_u^i.$$

Lemma 2.3. *Let $u \in \mathcal{A}$. Then the following statements hold:*

- 1) *there exists $Q(\xi) \in \mathcal{A}[\xi]$ with either $Q(\xi) = 0$ or $\deg Q(\xi) \leq d - 2$ such that*

$$(2.4) \quad \text{ad}_{-u} P(D) = \sum_{i=1}^n (D_i u)(\partial_i P)(D) + Q(D).$$

- 2) $(\text{ad}_{-u})^d P(D) = d! P(\nabla u).$

Proof: 1) First, if $\deg P(\xi) = 0$, then the statement holds trivially, for \mathcal{A} is commutative and hence $\text{ad}_{-u} P(\xi) = 0$. So we assume $\deg P(\xi) \geq 1$. By the linearity and also the commutativity of \mathcal{A} we may assume $P(\xi) = \xi_{i_1} \xi_{i_2} \cdots \xi_{i_k}$ with $1 \leq i_j \leq n$ for all $1 \leq j \leq k$.

We use the induction on $k \geq 1$. If $k = 1$, then $\text{ad}_{-u} D_{i_1} = \ell_{D_{i_1} u}$. Hence the statement holds by choosing $Q(\xi) = 0$. Assume that the statement holds for all $1 \leq k \leq m - 1$ and consider the case $k = m$.

Since ad_{-u} is a derivation of $\mathcal{W}(\mathcal{A})$, we have

$$\begin{aligned} \text{ad}_{-u} P(D) &= \sum_{j=1}^m D_{i_1} \cdots (\text{ad}_{-u} D_{i_j}) \cdots D_{i_m} = \sum_{j=1}^m D_{i_1} \cdots (\ell_{D_{i_j} u}) \cdots D_{i_m} \\ &= \sum_{j=1}^m (\ell_{D_{i_j} u}) D_{i_1} \cdots \widehat{D_{i_j}} \cdots D_{i_m} + \sum_{j=2}^m [D_{i_1} \cdots D_{i_{j-1}}, \ell_{D_{i_j} u}] D_{i_{j+1}} \cdots D_{i_m} \end{aligned}$$

Here $\widehat{D_{i_j}}$ means that the term D_{i_j} is omitted:

$$= \sum_{j=1}^m (D_{i_j} u) D_{i_1} \cdots \widehat{D_{i_j}} \cdots D_{i_m} + \sum_{j=2}^m (\text{ad}_{-D_{i_j} u} (D_{i_1} \cdots D_{i_{j-1}})) D_{i_{j+1}} \cdots D_{i_m}.$$

Applying the induction assumption to the terms $\text{ad}_{-D_{i_j} u} (D_{i_1} \cdots D_{i_{j-1}})$ ($2 \leq j \leq m$) in the sum above we see that there exists $Q(\xi) \in \mathcal{A}[\xi]$

with $Q(\xi) = 0$ or $\deg Q(\xi) \leq m - 2$ such that

$$\begin{aligned} \text{ad}_{-u} P(D) &= \sum_{j=1}^m (D_{i_j} u) D_{i_1} \cdots \widehat{D_{i_j}} \cdots D_{i_m} + Q(D) \\ &= \sum_{i=1}^n (D_i u) (\partial_i P)(D) + Q(D). \end{aligned}$$

Hence by the induction statement 1) follows.

2) First, by statement 1) it is easy to see that $(\text{ad}_{-u})^d P(D) = (\text{ad}_{-u})^d P_d(D)$. Then by the linearity and also the commutativity of \mathcal{A} we may assume $P_d(\xi) = \xi_{i_1} \xi_{i_2} \cdots \xi_{i_d}$ with $1 \leq i_j \leq n$ for all $1 \leq j \leq d$. Applying statement 1) (d times) we have

$$(\text{ad}_{-u})^d P(D) = \sum_{1 \leq k_1, k_2, \dots, k_d \leq n} (D_{k_1} u) (D_{k_2} u) \cdots (D_{k_d} u) (\partial_{k_1} \partial_{k_2} \cdots \partial_{k_d} P).$$

Then by the equation above and the commutativity of \mathcal{A} , it is easy to see that the equation in statement 2) follows. \square

Now we can prove the main result of this section.

Proof of Theorem 2.1: By Eq. (2.3) and Lemma 2.3, 2) we have

$$(2.5) \quad d! P_d(\nabla u) = (-1)^d \sum_{i=0}^d (-1)^i \binom{d}{i} u^{d-i} P(D) \circ \ell_u^i.$$

By applying both sides of the equation above to $1_{\mathcal{A}} \in \mathcal{A}$ and then using the condition $u^i \in \text{Ker } P(D)$ ($1 \leq i \leq d$), we get $d! P_d(\nabla u) = (-1)^d u^d P(D) \cdot 1_{\mathcal{A}}$. It is well-known and also easy to check that every derivation of a commutative ring annihilates the identity element of the ring. Hence $d! P_d(\nabla u) = (-1)^d u^d a_0$, i.e., Eq. (2.1) follows. Similarly, by applying Eq. (2.5) above to $u \in \mathcal{A}$ and using the condition $u^i \in \text{Ker } P(D)$ ($1 \leq i \leq d+1$), we get $d! P_d(\nabla u) u = 0$. Then by Eq. (2.1) we get Eq. (2.2). \square

One immediate consequence of Theorem 2.1 is the following

Corollary 2.4. *Let $D, P(\xi), a_0$ be as in Theorem 2.1, and $\text{nil}(\mathcal{A})$ the nil-radical of \mathcal{A} , i.e., the set of all nilpotent elements of \mathcal{A} . Then the following statements hold:*

- 1) $\mathfrak{r}(\text{Ker } \Lambda) \subseteq \text{Ann}(a_0)$, where $\text{Ann}(a_0)$ is the set of the elements $b \in \mathcal{A}$ such that $a_0 b = 0$;
- 2) if a_0 is not zero nor a zero-divisor of \mathcal{A} , then $\mathfrak{r}(\text{Ker } P(D)) = \text{nil}(\mathcal{A})$ and $\text{Ker } P(D)$ is a MS of \mathcal{A} ;

- 3) if $a_0 = 0$, then $\mathfrak{r}(\text{Ker } P(D)) \subseteq \{u \in \mathcal{A} \mid P_d(\nabla u) = 0\}$. In particular, if $n = 1$, i.e., D is a single derivation of \mathcal{A} , and the leading coefficient of $P(\xi)$ is not a zero-divisor of \mathcal{A} , then $\mathfrak{r}(\text{Ker } P(D)) \subseteq \{u \in \mathcal{A} \mid Du \in \text{nil}(\mathcal{A})\}$.

Example 2.5. Let $R = \mathbb{C}$ and \mathcal{A} the \mathbb{C} -algebra of all smooth complex valued functions $f(x)$ over \mathbb{R} . Let $D = \frac{d}{dx}$. Then for each nonzero univariate polynomial $P(\xi) \in \mathbb{C}[\xi]$, $\text{Ker } P(D)$ is the set of solutions $f(x) \in \mathcal{A}$ of the ordinary differential equation $P(D)f = 0$.

Let λ_i ($1 \leq i \leq k$) be the set of all distinct roots of $P(\xi)$ in \mathbb{C} with multiplicity m_i . Then it is well-known in the theory of ODE (e.g., see [L] or any other standard text book on ODE) that $\text{Ker } P(D)$ is the \mathbb{C} -subspace of \mathcal{A} spanned by $x^j e^{\lambda_i x}$ for all $1 \leq i \leq k$ and $1 \leq j \leq m_i$.

From the fact above it is easy to verify directly that $\mathfrak{r}(\text{Ker } P(D)) = \{0\}$, if $P(0) \neq 0$; and $\mathfrak{r}(\text{Ker } P(D)) = \mathbb{C}$, if $P(0) = 0$. Consequently, Theorem 2.1, Corollary 2.4 and also Proposition 5.4 in Section 5 all hold in this case.

We end this section with the following two remarks.

First, we will show in Propositions 3.7 and 5.4 that for the ordinary differential operators Φ of certain R -algebras \mathcal{A} (not necessarily commutative), the radical $\mathfrak{r}(\text{Ker } \Phi)$ also satisfies some other necessarily conditions (other than those in Theorem 2.1).

Second, Theorem 2.1 and Corollary 2.4 do not always hold for the differential operators of a noncommutative algebra, which can be seen from the following

Example 2.6. Let X, Y be two noncommutative free variables and $R[X, Y]$ the polynomial algebra in X and Y over R . Let J be the two-sided ideal of $R[X, Y]$ generated by Y^2 and $\mathcal{A} := R[X, Y]/J$. Let $D = \partial/\partial X$ and $P(\xi) = 1 - X\xi \in \mathcal{A}[\xi]$. Then $P(D) = I - \ell_X D$, where I denotes the identity map of \mathcal{A} , and ℓ_X the multiplication map by X from the left. Let $f = XY \in \mathcal{A}$. Then it is easy to check that for all $m \geq 1$, we have $P(D)(f^m) = 0$ but $P(D)(Xf^m) = -Xf^m \neq 0$. Therefore, $0 \neq f \in \mathfrak{r}(\text{Ker } P(D))$ and $\text{Ker } P(D)$ is not a MS of \mathcal{A} .

3. Some Cases for Non-Commutative Algebras

In this section, unless stated otherwise, R denotes a commutative ring such that the abelian group $(R, +)$ is torsion-free, and \mathcal{A} an R -algebra (not necessarily commutative) that is torsion-free as an R -module.

We denote by $I_{\mathcal{A}}$ or simply I the identity map of \mathcal{A} , and $\text{nil}(\mathcal{A})$ the set of all nilpotent elements of \mathcal{A} . We say \mathcal{A} is *reduced* if $\text{nil}(\mathcal{A}) = \{0\}$.

Furthermore, for each $a \in \mathcal{A}$, we denote by $\text{Ann}_\ell(a)$ the set of elements $b \in \mathcal{A}$ such that $ab = 0$.

Let D be an R -derivation of \mathcal{A} . We say that \mathcal{A} is *decomposable w.r.t.* (with respect to) the R -derivation D if \mathcal{A} can be written as a direct sum of the generalized eigen-subspaces of D . More precisely, let H be the set of all generalized eigenvalues of D in R and $\mathcal{A}_\lambda = \sum_{i=1}^{\infty} \text{Ker}(D - \lambda I)^i$ for each $\lambda \in H$. Then $\mathcal{A} = \bigoplus_{\lambda \in H} \mathcal{A}_\lambda$. It is easy to verify inductively that for all $m \geq 1$, $a, b \in \mathcal{A}$ and $\lambda, \mu \in R$, we have

$$(3.6) \quad (D - (\lambda + \mu)I)^m(ab) = \sum_{i=0}^m \binom{m}{i} ((D - \lambda I)^i a) ((D - \mu I)^{m-i} b).$$

Then by the identity above we have that $\mathcal{A}_\lambda \mathcal{A}_\mu \subseteq \mathcal{A}_{\lambda+\mu}$ for all $\lambda, \mu \in H$. In other words, the decomposition $\mathcal{A}_\lambda = \sum_{i=1}^{\infty} \text{Ker}(D - \lambda I)^i$ above is actually an additive R -algebra grading of \mathcal{A} .

Some examples of R -derivations with respect to which \mathcal{A} is decomposable are *semi-simple* R -derivations, for which \mathcal{A}_λ ($\lambda \in H$) coincides with the eigen-space of D corresponding to the eigenvalue λ of D , and also *locally finite* derivations when the base ring R is an algebraically closed field.

Once and for all, we let D_i ($1 \leq i \leq n$) be n commuting R -derivations of \mathcal{A} , i.e., $D_i D_j = D_j D_i$ for all $1 \leq i, j \leq n$, such that \mathcal{A} is decomposable w.r.t. each D_i . Then there exists a semi-subgroup Λ of the abelian group $(R^n, +)$ such that

$$(3.7) \quad \mathcal{A} = \bigoplus_{\lambda \in \Lambda} \mathcal{A}_\lambda,$$

where for each $\lambda = (k_1, k_2, \dots, k_n) \in \Lambda$,

$$(3.8) \quad \mathcal{A}_\lambda = \bigcap_{i=1}^n \left(\sum_{j=1}^{\infty} \text{Ker}(D_i - k_i I)^j \right).$$

In particular,

$$(3.9) \quad \mathcal{A}_0 = \bigcap_{i=1}^n \left(\sum_{j=1}^{\infty} \text{Ker } D_i^j \right).$$

Note also that each \mathcal{A}_λ ($\lambda \in \Lambda$) is invariant under D_i ($1 \leq i \leq n$), and $\mathcal{A}_\lambda \mathcal{A}_\mu \subseteq \mathcal{A}_{\lambda+\mu}$ for all $\lambda, \mu \in \Lambda$.

Now, let $\xi = (\xi_1, \xi_2, \dots, \xi_n)$ be n commutative free variables and $0 \neq P(\xi) \in R[\xi]$. We set $D := (D_1, D_2, \dots, D_n)$ and write $P(\xi) = \sum_{k=0}^d P_k(\xi)$ for some $d \geq 0$ and homogeneous polynomials $P_k(\xi)$ ($1 \leq k \leq d$) of degree k in ξ . We let $P(D)$ be the differential operator of \mathcal{A} obtained by replacing ξ_i by D_i ($1 \leq i \leq n$). Since D_i 's commute with one another, $P(D)$ is well-defined.

The first main result of this section is the following theorem which in some sense extends Theorem 2.1 to the differential operator $P(D)$ of the R -algebra \mathcal{A} which is not necessarily commutative.

Theorem 3.1. *With the setting as above, assume further that \mathcal{A} is reduced. Then the following statements hold:*

- 1) if $P_0 = P(0) = 0$ and $P_k(\xi)$ ($1 \leq k \leq d$) have no nonzero common zeros in R^n , then $\mathfrak{r}(\text{Ker } P(D)) \subseteq \mathcal{A}_0$;
- 2) if $P_0 = P(0) \neq 0$, then $\mathfrak{r}(\text{Ker } P(D)) = \{0\}$, and $\text{Ker } P(D)$ is a MS of \mathcal{A} .

In order to show the theorem above, we first need to show some lemmas.

Lemma 3.2. *Let R be an arbitrary commutative ring and \mathcal{A} an R -algebra that is torsion-free as an R -module. Let D and $P(\xi)$ be fixed as above. Then the following statements hold:*

- 1) $\text{Ker } P(D)$ is homogeneous w.r.t. the grading of \mathcal{A} in Eq. (3.7), i.e.,

$$(3.10) \quad \text{Ker } P(D) = \bigoplus_{\lambda \in \Lambda} (\mathcal{A}_\lambda \cap \text{Ker } P(D)).$$

- 2) Let $\mathcal{Z}_\Lambda(P)$ be the set of $\lambda \in \Lambda$ such that $P(\lambda) = 0$. Then

$$(3.11) \quad \text{Ker } P(D) \subseteq \bigoplus_{\lambda \in \mathcal{Z}_\Lambda(P)} \mathcal{A}_\lambda.$$

Proof: 1) Since for each $\lambda \in \Lambda$, \mathcal{A}_λ is preserved by D_i ($1 \leq i \leq n$), and hence also is preserved by $P(D)$, from which Eq. (3.10) follows.

2) Let $0 \neq u \in \mathcal{A}$ and write $u = \sum_{i=1}^\ell u_{\lambda_i}$ for some distinct $\lambda_i \in \Lambda$ ($1 \leq i \leq n$) and $u_{\lambda_i} \in \mathcal{A}_{\lambda_i}$. Then by Eq. (3.10) we have that $u \in \text{Ker } P(D)$, if and only if $u_{\lambda_i} \in \text{Ker } P(D)$. So we may assume $\ell = 1$ and $u \in \mathcal{A}_\lambda$ for some $\lambda \in \Lambda$.

Write $\lambda = (k_1, k_2, \dots, k_n)$. For each $1 \leq j \leq n$, we define a non-negative integer r_j as follows.

First, let r_n be the greatest non-negative integer such that $(D_n - k_n I)^{r_n} u \neq 0$, and inductively, for each $1 \leq j \leq n-1$, let r_j be the greatest non-negative integer such that $(D_j - k_j I)^{r_j} \left(\prod_{s=j+1}^n (D_s - k_s I)^{r_s} \right) u \neq 0$.

Set $\tilde{u} := \left(\prod_{j=1}^n (D_j - k_j I)^{r_j} \right) u$. Then $0 \neq \tilde{u} \in \mathcal{A}_\lambda$, $\tilde{u} \in \text{Ker } P(D)$, and $D_j \tilde{u} = k_j \tilde{u}$ for all $1 \leq j \leq n$. Hence $0 = P(D) \tilde{u} = P(\lambda) \tilde{u}$. Since \mathcal{A} is torsion-free as an R -module, we have $P(\lambda) = 0$, as desired. \square

Definition 3.3. Let A be a subset of R^n and $\lambda \in A$. We say λ is an extremal element of A if for all $m \geq 1$, $m\lambda$ can not be written as a linear combination of other elements of A with positive integer coefficients whose sum is less or equal to m .

The following lemma should be well-known. But for the sake of completeness, we here include a direct proof.

Lemma 3.4. Let R be a commutative ring such that the abelian group $(R, +)$ is torsion free. Then every nonempty finite subset A of R^n has at least one extremal element.

Proof: Write $A = \{\lambda_1, \lambda_2, \dots, \lambda_n\}$ with $\lambda_i \neq \lambda_j$ for all $1 \leq i \neq j \leq n$. We use the induction on n . If $n = 1$, there is nothing to show. So we assume $n \geq 2$.

Consider first the case $n = 2$ with $\lambda_2 \neq 0$. If the lemma fails, then $m_1\lambda_1 = k_1\lambda_2$ and $m_2\lambda_2 = k_2\lambda_1$ for some $m_i, k_i \geq 1$ with $k_i \leq m_i$. Then $m_1m_2\lambda_2 = m_1(k_2\lambda_1) = k_2(m_1\lambda_1) = k_1k_2\lambda_2$. Hence $m_1m_2 = k_1k_2$, for $\lambda_2 \neq 0$ and $(R, +)$ is torsion-free, from which we have $m_1 = k_1$ (and $m_2 = k_2$). By the fact that $(R, +)$ is torsion-free again, we have $\lambda_1 = \lambda_2$. Contradiction.

Now assume the lemma holds for all $2 \leq n \leq k$ and consider the case $n = k + 1$. If λ_{k+1} is an extremal point of A , then there is nothing to show. Assume otherwise. Then there exist $m \geq 1$ and $c_i \in \mathbb{N}$ ($1 \leq i \leq k$) such that

$$(3.12) \quad m\lambda_{k+1} = \sum_{i=1}^k c_i\lambda_i,$$

$$(3.13) \quad 1 \leq \sum_{i=1}^k c_i \leq m.$$

By the induction assumption the set $A' := \{\lambda_1, \lambda_2, \dots, \lambda_k\}$ has an extremal element, say, λ_1 . We claim that λ_1 is also an extremal point of the set A . Otherwise, there exist $q \geq 1$ and $c'_j \in \mathbb{N}$ ($2 \leq j \leq k + 1$) such that

$$(3.14) \quad q\lambda_1 = c'_{k+1}\lambda_{k+1} + \sum_{j=2}^k c'_j\lambda_j,$$

$$(3.15) \quad 1 \leq c'_{k+1} + \sum_{j=2}^k c'_j \leq q.$$

Then by Eqs. (3.14) and (3.12) we have

$$\begin{aligned}
 (3.16) \quad mq\lambda_1 &= mc'_{k+1}\lambda_{k+1} + m \sum_{j=1}^k c'_j \lambda_j \\
 &= c'_{k+1} \sum_{i=1}^k c_i \lambda_i + m \sum_{j=1}^k c'_j \lambda_j.
 \end{aligned}$$

For the sum of all the coefficients of the linear combination on the right hand side of the equation above, by Eqs. (3.13) and (3.15) we have

$$\begin{aligned}
 (3.17) \quad 1 &\leq c'_{k+1} \sum_{i=1}^k c_i + m \sum_{j=1}^k c'_j \leq c'_{k+1}m + m \sum_{j=1}^k c'_j \\
 &= m(c'_{k+1} + \sum_{j=1}^k c'_j) \leq mq.
 \end{aligned}$$

Then by Eqs. (3.16) and (3.17), λ_1 is not an extremal element of A' , which contradicts to the choice of λ_1 . Therefore λ_1 is an extremal point of A , and the lemma follows. \square

Lemma 3.5. *Let $0 \neq u \in \mathfrak{r}(\text{Ker } P(D))$ and write $u = \sum_{i=1}^\ell u_{\lambda_i}$ for some distinct $\lambda_i \in \Lambda$ ($1 \leq i \leq \ell$) and $0 \neq u_{\lambda_i} \in \mathcal{A}_{\lambda_i}$. Then for each extremal element λ_j of the set $\{\lambda_i \mid 1 \leq i \leq \ell\}$, either u_{λ_j} is nilpotent, or $P_k(\lambda_j) = 0$ for all $0 \leq k \leq d$.*

Proof: Assume that u_{λ_j} is not nilpotent. Since λ_j is an extremal element of the set $\{\lambda_i \mid 1 \leq i \leq \ell\}$, it is easy to see that for each $m \geq 1$, the homogeneous component of u^m in $\mathcal{A}_{m\lambda_j}$ is equal to $u_{\lambda_j}^m$. Since $u^m \in \text{Ker } P(D)$ when $m \gg 0$, by Lemma 3.2, 1) and 2) we have $u_{\lambda_j}^m \in \text{Ker } P(D)$ and $P(m\lambda_j) = 0$ for all $m \gg 0$. More explicitly, for all $m \gg 0$, we have

$$0 = P(m\lambda_j) = \sum_{k=0}^d m^k P_k(\lambda_j).$$

Since $(R, +)$ is torsion-free, by the vandemonde determinant we have $P_k(\lambda_j) = 0$ for all $0 \leq k \leq d$. \square

Proof of Theorem 3.1: Let $0 \neq u \in \mathfrak{r}(\text{Ker } P(D))$ and write $u = \sum_{i=1}^\ell u_{\lambda_i}$ for some distinct $\lambda_i \in \Lambda$ ($1 \leq i \leq \ell$) and $0 \neq u_{\lambda_i} \in \mathcal{A}_{\lambda_i}$. Let B be the set of all nonzero λ_i ($1 \leq i \leq \ell$). If $B \neq \emptyset$, then by Lemma 3.4, B has at least one extremal element, say λ_j . Then by Definition 3.3,

λ_j is also an extremal element of the set $B \cup \{0\}$. Since \mathcal{A} is reduced, u_{λ_j} is not nilpotent. Then by Lemma 3.5, $P_k(\lambda_j) = 0$ for all $0 \leq k \leq d$.

If $P_0 = 0$, and $P_k(\xi)$ ($1 \leq k \leq d$) have no nonzero common zero in R^n , then we have $\lambda_j = 0$, which is a contradiction. Therefore in this case $B = \emptyset$ and $u \in \mathcal{A}_0$, whence the statement 1) follows.

If $P_0 \neq 0$, then we also have $B = \emptyset$ and $u \in \mathcal{A}_0$, for $P_0(\lambda_j) = P_0 \neq 0$. Furthermore, since $P(0) = P_0 \neq 0$, by Lemma 3.2, 2) we have $\mathcal{A}_0 \cap \text{Ker } P(D) = 0$, whence $u = 0$. Contradiction. Therefore statement 2) follows. \square

Next, we show that Theorem 3.1 with some extra conditions also holds for commuting locally finite R -derivations. Recall that an R -derivation δ of an R -algebra \mathcal{A} is *locally finite (over R)* if for each $u \in \mathcal{A}$, the R -submodule of \mathcal{A} spanned by elements $\delta^k u$ ($k \geq 0$) over R is finitely generated as an R -module.

Proposition 3.6. *Assume that R is an integral domain of characteristic zero and \mathcal{A} is a reduced R -algebra that is torsion-free as an R -module. Denote by K_R the field of fractions of R and \bar{K}_R the algebraic closure of K_R . Let $P(\xi) \in R[\xi]$ and $D = (D_1, D_2, \dots, D_n)$ be n commuting locally finite R -derivations of \mathcal{A} . Write $P(\xi) = \sum_{i=0}^d P_i(\xi)$ with $P_i(\xi)$ ($0 \leq i \leq d$) being homogeneous of degree i . Then the following statements hold:*

- 1) *if $P(0) = 0$ and $P_i(\xi)$ ($1 \leq i \leq d$) have no nonzero common zeros in \bar{K}_R^n , then we have $\mathfrak{r}(\text{Ker } P(D)) \subseteq \mathcal{A}_0$, where $\mathcal{A}_0 = \bigcap_{i=1}^n (\sum_{m=1}^{\infty} \text{Ker } D_i^m)$;*
- 2) *if $P(0) \neq 0$, then $\mathfrak{r}(\text{Ker } P(D)) = \{0\}$, and $\text{Ker } P(D)$ is a MS of \mathcal{A} .*

Proof: Set $\bar{\mathcal{A}} = \bar{K}_R \otimes_R \mathcal{A}$. Since \mathcal{A} is torsion-free as an R -module, the standard map $\mathcal{A} \simeq R \otimes_R \mathcal{A} \rightarrow K_R \otimes_R \mathcal{A}$ is injective, for by [AM, Prop. 3.3] $K_R \otimes_R \mathcal{A}$ is isomorphic to the localization $S^{-1}\mathcal{A}$ with $S = R \setminus \{0\}$. Since every field is absolutely flat, the standard map $K_R \otimes_R \mathcal{A} \rightarrow \bar{K}_R \otimes_R \mathcal{A}$ is also injective. Therefore, we may view \mathcal{A} as an R -subalgebra in the standard way and extend D \bar{K}_R -linearly to $\bar{\mathcal{A}}$, which we denote by $\bar{D} = (\bar{D}_1, \bar{D}_2, \dots, \bar{D}_n)$.

Note that \bar{D}_i ($1 \leq i \leq n$) are n commuting \bar{K}_R -derivations of $\bar{\mathcal{A}}$, which are also locally finite over \bar{K}_R . Then $\bar{\mathcal{A}}$ by [E1, Proposition 1.3.8]) is decomposable w.r.t. \bar{D}_i for each $1 \leq i \leq n$. By applying Theorem 3.1 to $P(\bar{D})$ and using the fact $\bar{\mathcal{A}}_0 \cap \mathcal{A} = \mathcal{A}_0$ we see that the proposition follows. \square

Next, we use the proposition above to show that Corollary 2.4 with some extra conditions can be extended to the ordinary differential operators of some noncommutative algebras.

Proposition 3.7. *Let R, \mathcal{A} be as in Proposition 3.6 and D an arbitrary (single) R -derivation of \mathcal{A} . Then for every univariate polynomial in $0 \neq P(\xi) \in R[\xi]$, the following statements hold:*

- 1) *if $P(0) = 0$, then $\mathfrak{r}(\text{Ker } P(D)) \subseteq \mathfrak{r}(\mathcal{A}_0)$, where $\mathcal{A}_0 = \sum_{i=1}^{\infty} \text{Ker } D^i$;*
- 2) *if $P(0) \neq 0$, then $\mathfrak{r}(\text{Ker } P(D)) = \{0\}$, and $\text{Ker } P(D)$ is a MS of \mathcal{A} .*

Proof: The case $\deg P(\xi) = 0$ is trivial. So we assume $\deg P(\xi) \geq 1$. Let K_R be the field of fractions of R with the algebraic closure \bar{K}_R , and set $\bar{\mathcal{A}} = \bar{K}_R \otimes_R \mathcal{A}$. As pointed out in the proof of Proposition 3.6 we may view \mathcal{A} as an R -subalgebra of $\bar{\mathcal{A}}$ in the standard way and extend D \bar{K}_R -linearly to $\bar{\mathcal{A}}$, which we denote by \bar{D} .

Let $V = \text{Ker } P(D)$. Then V is an R -subspace of \mathcal{A} preserved by D . Set $\bar{V} = \bar{K}_R \otimes_R V$. Then $\bar{D}|_{\bar{V}}$ as a \bar{K}_R -linear map from \bar{V} to \bar{V} is algebraic over \bar{K}_R , for $P(D|_V) = P(D)|_V = 0$ and hence $P(\bar{D}|_{\bar{V}}) = 0$. It is well-known (e.g., see [H, Proposition 4.2]) that \bar{V} can be decomposed as a direct sum of the generalized eigen-spaces of $\bar{D}|_{\bar{V}}$. Let $\bar{\mathcal{B}}$ be the \bar{K}_R -subalgebra of $\bar{\mathcal{A}}$ generated by elements of \bar{V} . Then $\bar{\mathcal{B}}$ is \bar{D} -invariant. Furthermore, by Eq. (3.6) it is easy to see that $\bar{\mathcal{B}}$ is decomposable w.r.t. $\bar{D}|_{\bar{\mathcal{B}}}$.

Now let $u \in \mathfrak{r}(\text{Ker } P(D))$. Then there exists $N \geq 1$ such that $u^m \in \text{Ker } P(D)$, and hence is also in $\bar{\mathcal{B}}$, for all $m \geq N$. Consequently, we also have $u^m \in \mathfrak{r}(\text{Ker } P(\bar{D}|_{\bar{\mathcal{B}}}))$ for all $m \geq N$. Note that $P_k(\xi)$ ($1 \leq k \leq d$) have no nonzero common zero in \bar{K}_R , for $P(\xi)$ is a univariate polynomial of degree greater or equal to 1. Therefore, if $P(0) = 0$, then by applying Proposition 3.6, 1) to $P(\bar{D}|_{\bar{\mathcal{B}}})$ (as a differential operator of $\bar{\mathcal{B}}$), we have $u^m \in \sum_{i=1}^{\infty} \text{Ker } \bar{D}^i$ for all $m \geq N$. Since $\text{Ker } D^i = \mathcal{A} \cap \text{Ker } \bar{D}^i$ for all $i \geq 1$, we further have $u^m \in \mathcal{A}_0 = \sum_{i=1}^{\infty} D^i$ for all $m \geq N$. Hence $u \in \mathfrak{r}(\mathcal{A}_0)$ and statement 1) follows.

If $P(0) \neq 0$, then by applying Proposition 3.6, 2) to $P(\bar{D}|_{\bar{\mathcal{B}}})$ (as a differential operator of $\bar{\mathcal{B}}$), we have $u^m = 0$ for all $m \geq N$. Since \mathcal{A} is reduced, we have $u = 0$. Hence statement 2) also follows. \square

We end this section with the following open problem which is worthy of further investigations.

Open Problem 3.8. *Let R be an arbitrary commutative ring and \mathcal{A} an arbitrary unital noncommutative R -algebra. Let $D = (D_1, D_2, \dots, D_n)$ be n R -derivations of \mathcal{A} , and $Q(\xi) \in R[\xi]$ a polynomial in n noncommutative free variables $\xi = (\xi_1, \xi_2, \dots, \xi_n)$. Set $a_0 := Q(0)$ and denote*

by $\text{Ann}_\ell(a_0)$ the set of all elements $b \in \mathcal{A}$ such that $a_0b = 0$. Decide whether or not it is always true that $\mathfrak{r}(\text{Ker } Q(D)) \subseteq \mathfrak{r}(\text{Ann}_\ell(a_0))$?

4. Some Applications to Locally Algebraic Derivations

In this section we use some results proved in the last two sections to derive some properties of locally algebraic or locally integral derivations.

Definition 4.1. Let R be a unital commutative ring, \mathcal{A} an R -algebra and D an R -derivation of \mathcal{A} .

- 1) We say D is algebraic over R if there exists a nonzero polynomial $p(t) \in R[t]$ such that $p(D) = 0$.
- 2) We say D is locally algebraic over R if for each $a \in \mathcal{A}$, there exists a D -invariant R -subalgebra \mathcal{A}_1 of \mathcal{A} containing a , and a nonzero polynomial $p_a(t) \in R[t]$ such that $p_a(D)|_{\mathcal{A}_1} = 0$.

If $p(t)$ in statement 1) (resp., $p_a(t)$ in statement 2) for all $a \in \mathcal{A}$) of the definition above can be chosen to be a monic polynomial, we say D is integral (resp., locally integral) over R .

An example of a derivation that is locally algebraic but not algebraic is as follows.

Example 4.2. Let x_i ($i \geq 1$) be a sequence of free commutative variables and $\mathbb{C}[x_i | i \geq 1]$ the polynomial algebra over \mathbb{C} in x_i ($i \geq 1$). Let I be the ideal generated by x_i^{i+1} ($i \geq 1$) and $\mathcal{A} = \mathbb{C}[x_i | i \geq 1]/I$. Then it can be readily verified that $D := \sum_{i=1}^{\infty} x_i \partial / \partial x_i$ is a well-defined \mathbb{C} -derivation of \mathcal{A} , which is locally algebraic but not (globally) algebraic over \mathbb{C} .

Theorem 4.3. Let R be a commutative ring and \mathcal{A} a commutative R -algebra such that the abelian group $(\mathcal{A}, +)$ is torsion-free. Then for every R -derivation D of \mathcal{A} that is locally integral over R , the image $\text{Im } D := D(\mathcal{A}) \subseteq \text{nil}(\mathcal{A})$, where $\text{nil}(\mathcal{A})$ denotes the nil-radical of \mathcal{A} , i.e., the set of nilpotent elements of \mathcal{A} .

Proof: Let $a \in \mathcal{A}$, and \mathcal{A}_1 be a D -invariant R -subalgebra of \mathcal{A} and $p_a(t)$ a monic polynomial in $R[t]$ such that $a \in \mathcal{A}_1$ and $p_a(D)|_{\mathcal{A}_1} = 0$. Then $\mathcal{A}_1 \subseteq \text{Ker } p_a(D)$. In particular, $a^m \in \text{Ker } p_a(D)$ for all $m \geq 1$.

Replacing $p_a(t)$ by $tp_a(t)$ we assume $p_a(0) = 0$ and $d := \deg p_a(t) \geq 1$. Then by Theorem 2.1 we have $d!(Da)^d = 0$. Since $(\mathcal{A}, +)$ as an abelian group is torsion-free, we have $(Da)^d = 0$, whence $Da \in \text{nil}(\mathcal{A})$ and the theorem follows. \square

Since every nilpotent R -derivation of \mathcal{A} is locally integrable over R , by Theorem 4.3 we immediately have the following

Corollary 4.4. *Let R, \mathcal{A} be as in Theorem 4.3 and D a nilpotent R -derivation of \mathcal{A} . Then $\text{Im } D \subseteq \text{nil}(\mathcal{A})$.*

Furthermore, from the proof of Theorem 4.3 it is also easy to see that we have the following

Corollary 4.5. *Let R and \mathcal{A} be as in Theorem 4.3. Assume further that \mathcal{A} is torsion-free as an R -module. Then for every R -derivation D of \mathcal{A} that is locally algebraic over R , we have $\text{Im } D \subseteq \text{nil}(\mathcal{A})$.*

Next, we consider the R -derivations of some reduced R -algebra \mathcal{A} (not necessarily commutative) that are locally algebraic over R .

Theorem 4.6. *Let R be a unital integral domain of characteristic zero and \mathcal{A} a unital reduced R -algebra (not necessarily commutative) that is torsion-free as an R -module. Then \mathcal{A} has no nonzero R -derivations that are locally algebraic over R . In particular, \mathcal{A} has no nonzero nilpotent R -derivations.*

Proof: Let D be an R -derivation of \mathcal{A} that is locally algebraic over R . Let $a \in \mathcal{A}$, and \mathcal{A}_1 be a D -invariant R -subalgebra of \mathcal{A} and $0 \neq p_a(t) \in R[t]$ such that $a \in \mathcal{A}_1$ and $p_a(D)|_{\mathcal{A}_1} = 0$. Then $a^m \in \mathcal{A}_1 \subseteq \text{Ker } p_a(D)$ for all $m \geq 1$, whence $a \in \mathfrak{r}(\text{Ker } p_a(D))$.

Replacing $p_a(t)$ by $tp_a(t)$ we assume $p_a(0) = 0$. Then by applying Proposition 3.7, 1) to the differential operator $p_a(D)$, we have $a \in \mathfrak{r}(\mathcal{A}_0)$, where $\mathcal{A}_0 = \sum_{i=1}^{\infty} \text{Ker } D^i$. Consequently, $\mathfrak{r}(\mathcal{A}_0) = \mathcal{A}$. Then by [Z3, Lemma 2.4] we have $\mathcal{A}_0 = \mathcal{A}$, i.e., D is locally nilpotent.

Let K_R be the field of fractions of R and $\mathcal{B} := K_R \otimes_R \mathcal{A}$. As pointed out in the proof of Proposition 3.6, we may view \mathcal{A} as an R -subalgebra of \mathcal{B} in the standard way and extend D K_R -linearly to \mathcal{B} , which we denote by \bar{D} .

Let $a, p_a(t)$ be fixed as above, and $N \geq 1$ such that $D^N a = 0$. Write $p_a(t) = t^k h(t)$ for some $k \geq 1$ and $h(t) \in K_R[t]$ with $h(0) \neq 0$. Then $p_a(\bar{D})a = 0$ and $\bar{D}^N a = 0$. Since $\gcd(p_a(t), t^N) = t^\ell$ in $K_R[t]$ with $\ell = \min\{k, N\}$, we have $\bar{D}^\ell a = 0$. Hence $D^k a = \bar{D}^k a = 0$ for all $a \in \mathcal{A}$. Therefore D is nilpotent. Then by [Z4, Lemma 6.1] we have $D = 0$, whence the theorem follows. \square

One remark on Theorem 4.6 is that without the characteristic zero condition, the theorem may be false, which can be seen from the following example. For more integral derivations of algebras over a field of characteristic $p > 0$, see [N].

Example 4.7. *Let K be a field of characteristic $p > 0$, $\mathcal{A} = K[x]$ and $D = d/dx$. Then $D^p = 0$. Hence D is a nonzero K -derivation of \mathcal{A} that is algebraic over K .*

One immediate consequence of Theorem 4.3, Corollary 4.5 and Theorem 4.6 is the following corollary which in some sense gives an affirm answer to the so-called LNED conjecture proposed in [Z4] for nilpotent, or locally integral, or locally algebraic derivations of certain algebras.

Corollary 4.8. 1) *Let R, \mathcal{A} be as in Theorem 4.3 and D an R -derivation of \mathcal{A} that is locally integral over R . Then D maps every R -subspace of \mathcal{A} to a MS of \mathcal{A} .*

2) *Let R, \mathcal{A} be as in Corollary 4.5 or as in Theorem 4.3, and D an R -derivation of \mathcal{A} that is locally algebraic over R . Then D maps every R -subspace of \mathcal{A} to a MS of \mathcal{A} .*

We end this section with the following

Proposition 4.9. *Let R be a commutative ring and \mathcal{A} a reduced R -algebra (not necessarily commutative) such that $(\mathcal{A}, +)$ is torsion-free. Let $r \geq 1$, $a \in \mathcal{A}$ and D be an R -derivation of \mathcal{A} such that $D^r a^m = 0$ for all $1 \leq m \leq 2^{r-1}$. Then $a \in \text{Ker } D$. Consequently, we have $\text{Ker } D \subseteq \mathfrak{r}(\text{Ker } D^k) = \mathfrak{r}(\text{Ker } D)$.*

Note that when \mathcal{A} is commutative, the lemma follows easily from Theorem 2.1. Here we give a proof independent on the commutativity of \mathcal{A} .

Proof of Proposition 4.9: The case $r = 1$ is obvious. So assume $r \geq 2$. Then $2r - 2 \geq r$ and for each $1 \leq k \leq 2^{r-2}$, by the Leibniz rule we have

$$0 = D^{2r-2} a^{2k} = \sum_{i=0}^{2r-2} \binom{2r-2}{i} (D^i a^k) (D^{2r-2-i} a^k)$$

Since $D^i a^k = 0$ for all $i \geq r$, there is only one term in the sum above that is not equal to 0, namely, the term with $i = r - 1$. Therefore $\binom{2r-2}{r-1} (D^{r-1} a^k)^2 = 0$. Since \mathcal{A} is reduced and $(\mathcal{A}, +)$ is torsion-free, we have $D^{r-1} a^k = 0$ for all $1 \leq k \leq 2^{(r-1)-1}$. Repeating the procedure above we see that $Da = 0$, i.e., $a \in \text{Ker } D$.

Now let $u \in \mathfrak{r}(\text{Ker } D^r)$. Then there exists $N \geq 1$ such that $u^m \in \text{Ker } D^r$ for all $m \geq N$. Applying the result shown above to u^m ($m \geq N$) we have $u^m \in \text{Ker } D$ for all $m \geq N$. Hence $u \in \mathfrak{r}(\text{Ker } D)$, and $\mathfrak{r}(\text{Ker } D^r) \subseteq \mathfrak{r}(\text{Ker } D)$. Conversely, since $\text{Ker } D$ is an R -subalgebra of \mathcal{A} and $\text{Ker } D \subseteq \text{Ker } D^r$, we also have $\text{Ker } D \subseteq \mathfrak{r}(\text{Ker } D^r)$ and $\mathfrak{r}(\text{Ker } D) \subseteq \mathfrak{r}(\text{Ker } D^r)$. Hence the proposition follows. \square

5. A Differential Vandemonde Determinant

Throughout this section \mathcal{A} stands for a commutative ring and D for a derivation of \mathcal{A} .

Proposition 5.1. *Let \mathcal{A} and D be fixed as above. Then for all $f \in \mathcal{A}$ and $n \geq 1$, we have*

$$(5.1) \quad \det \begin{pmatrix} f & f^2 & \cdots & f^n \\ D(f) & D(f^2) & \cdots & D(f^n) \\ D^2(f) & D^2(f^2) & \cdots & D^2(f^n) \\ \vdots & \vdots & \ddots & \vdots \\ D^{n-1}(f) & D^{n-1}(f^2) & \cdots & D^{n-1}(f^n) \end{pmatrix} = \alpha_n (Df)^{\frac{1}{2}n(n-1)} f^n,$$

where $\alpha_n = \prod_{k=1}^{n-1} k!$.

The idea of the proof is to show that the matrix in Eq. (5.1) can be transformed by some elementary column operations to an upper triangular matrix whose the $(i, i)^{th}$ diagonal entry is equal to $(i-1)!(Df)^{i-1}f$ for all $1 \leq i \leq n$. For example, for the case $n = 2$, by subtracting from the second column the multiple of the first column by f we get

$$(5.2) \quad \begin{pmatrix} f & f^2 \\ D(f) & D(f^2) \end{pmatrix} \Rightarrow \begin{pmatrix} f & 0 \\ D(f) & fD(f) \end{pmatrix}.$$

To see this can be achieved for all $n \geq 2$, it suffices to show the following lemma, from which Proposition 5.1 immediately follows.

Lemma 5.2. *Let D and f as in Proposition 5.1 and $k \geq 2$. Then there exist $\alpha_{k,j} \in \mathcal{A}$ ($1 \leq j \leq k-1$) such that for each $0 \leq i \leq k-1$, we have*

$$(5.3) \quad D^i(f^k) - \sum_{j=1}^{k-1} \alpha_{k,j} f^{k-j} D^i(f^j) = \delta_{i,k-1} (k-1)! (Df)^{k-1} f,$$

where $\delta_{i,k-1}$ is the Kronecker delta function.

Proof: We use induction on k . If $k = 2$, then $\alpha_{2,1} = 1$ solves the equations in Eq. (5.3), as already pointed out in Eq. (5.2) above.

Assume that lemma holds for some $k \geq 2$ and consider the case $k+1$. By writing f^{k+1} as $f \cdot f^k$ and applying the Leibniz rule, we have for each $0 \leq i \leq k$

$$D^i(f^{k+1}) - f D^i(f^k) = \sum_{\ell=0}^{i-1} \binom{i}{\ell} (D^{i-\ell} f) D^\ell(f^k)$$

Applying the induction assumption to $D^\ell(f^k)$ and noticing that $\ell = k - 1$, if and only if $i = k = \ell + 1$, since $\ell \leq i - 1 \leq k - 1$:

$$\begin{aligned}
&= \delta_{i,k} k(k-1)! (Df)^k f + \sum_{\ell=0}^{i-1} \binom{i}{\ell} (D^{i-\ell} f) \left(\sum_{j=1}^{k-1} \alpha_{k,j} f^{k-j} D^\ell(f^j) \right) \\
&= \delta_{i,k} k! (Df)^k f + \sum_{j=1}^{k-1} \alpha_{k,j} f^{k-j} \sum_{\ell=0}^{i-1} \binom{i}{\ell} (D^{i-\ell} f) D^\ell(f^j) \\
&= \delta_{i,k} k! (Df)^k f + \sum_{j=1}^{k-1} \alpha_{k,j} f^{k-j} \left(D^i(f^{j+1}) - f D^i(f^j) \right) \\
&= \delta_{i,k} k! (Df)^k f + \alpha_{k,k-1} f D^i(f^k) - \alpha_{k,1} f^k D(f) \\
&\quad + \sum_{j=2}^{k-1} (\alpha_{k,j-1} - \alpha_{k,j}) f^{k+1-j} D^i(f^j).
\end{aligned}$$

Set

$$(5.4) \quad \alpha_{k+1,j} := \begin{cases} -\alpha_{k,1} & \text{if } j = 1; \\ \alpha_{k,j-1} - \alpha_{k,j} & \text{if } 2 \leq j \leq k-1; \\ 1 + \alpha_{k,k-1} & \text{if } j = k. \end{cases}$$

Then $\alpha_{k+1,j}$ ($1 \leq j \leq k$) solve the equations in Eq. (5.3) for the case $k+1$ and $0 \leq i \leq k$. Hence, by induction the lemma follows. \square

Remark 5.3. One application of the formula in Eq. (5.1) is as follows. We first apply the formula to some special function $f(x)$ and derivation D , and then evaluate x at a fixed point c . By doing so, we may get formulas for the determinants of several families of matrices, e.g., letting $f = x^d$, $D = x^m \frac{d}{dx}$ and $c = \pm 1$ for all $d, m \in \mathbb{Z}$. In particular, if we choose $d = -1$, $m = 0$ and $c = 1$, then with a little more argument we get the following formula with $0! := 1$ for all $n \geq 1$:

$$(5.5) \quad \det \left((i+j-2)! \right)_{1 \leq i,j \leq n} = \left(\prod_{k=1}^{n-1} k! \right)^2.$$

Another consequence of Proposition 5.1 is the following

Proposition 5.4. Let D be a derivation of \mathcal{A} , ξ be a free variable and $0 \neq P(\xi) = \sum_{i=0}^d c_i \xi^i \in \mathcal{A}[\xi]$. Let $f \in \mathcal{A}$ such that $f^m \in \text{Ker } P(D)$ for all $1 \leq m \leq d+1$. Then for each $0 \leq i \leq d$, we have

$$(5.6) \quad \alpha_{d+1} c_i (Df)^{\frac{1}{2}(d+1)} f^{d+1} = 0,$$

where $\alpha_{d+1} = \prod_{k=1}^d k!$

Proof: Let B be the transpose of the matrix in Eq. (5.1) with $n-1 = d$. Since $P(D)(f^m) = 0$ for all $1 \leq m \leq d+1$, we have $Bv = 0$, where v denotes the column vector $(c_0, c_1, \dots, c_d)^\tau$. Then from Eq. (5.1) the proposition follows. \square

Corollary 5.5. *Let $D, f, P(\xi)$ be as in Proposition 5.4. Assume further that $(\mathcal{A}, +)$ is torsion-free and for some $0 \leq i \leq d$, c_i is not zero nor a zero-divisor of \mathcal{A} . Then fDf is nilpotent.*

Proof: By Proposition 5.4 we have $\alpha_{d+1}c_i(Df)^{\frac{1}{2}d(d+1)}f^{d+1} = 0$. Hence, $f^{d+1}(Df)^{\frac{1}{2}d(d+1)} = 0$, for $(\mathcal{A}, +)$ is torsion-free and c_i is not zero nor a zero-divisor of \mathcal{A} . Then $(fDf)^m = 0$ for all $m \geq \max\{d+1, \frac{1}{2}d(d+1)\}$, whence the corollary follows. \square

REFERENCES

- [AM] M. F. Atiyah and I. G. Macdonald, *Introduction to Commutative Algebra*. Addison-Wesley Publishing Co., 1969. [MR0242802].
- [BCW] H. Bass, E. Connell and D. Wright, *The Jacobian Conjecture, Reduction of Degree and Formal Expansion of the Inverse*. Bull. Amer. Math. Soc. **7**, (1982), 287–330.
- [DEZ] H. Derksen, A. van den Essen and W. Zhao, *The Gaussian Moments Conjecture and the Jacobian Conjecture*. To appear in *Israel J. Math.*. See also arXiv:1506.05192 [math.AC].
- [E1] A. van den Essen, *Polynomial Automorphisms and the Jacobian Conjecture*. Prog. Math., Vol.190, Birkhäuser Verlag, Basel, 2000.
- [E2] A. van den Essen, *Introduction to Mathieu Subspaces*. “International Short-School/Conference on Affine Algebraic Geometry and the Jacobian Conjecture” at Chern Institute of Mathematics, Nankai University, Tianjin, China. July 14-25, 2014.
- [EH] A. van den Essen and L. C. van Hove, *Mathieu-Zhao Spaces*. To appear.
- [EN] A. van den Essen and S. Nieman, *Mathieu-Zhao Spaces of Univariate Polynomial Rings with Non-zero Strong Radical*. J. Pure Appl. Algebra, **220** (2016), no. 9, 3300–3306.
- [EZ] A. van den Essen and W. Zhao, *Mathieu Subspaces of Univariate Polynomial Algebras*. J. Pure Appl. Algebra. 217 (2013), no.7, 1316–1324. See also arXiv:1012.2017[math.AC].
- [EWZ] A. van den Essen, D. Wright and W. Zhao, *Images of Locally Finite Derivations of Polynomial Algebras in Two Variables*. J. Pure Appl. Algebra **215** (2011), no.9, 2130–2134. [MR2786603]. See also arXiv:1004.0521[math.AC].
- [H] J. E. Humphreys, (1972), *Introduction to Lie Algebras and Representation Theory*. Graduate Texts in Mathematics, Springer, 1972.
- [Ke] O. H. Keller, *Ganze Gremona-Transformationen*. Monats. Math. Physik **47** (1939), no. 1, 299–306.

- [L] S. Lefschetz, *Differential Equations: Geometric Theory*. Reprinting of the second edition. Dover Publications, Inc., New York, 1977. [MR0435481].
- [M] O. Mathieu, *Some Conjectures about Invariant Theory and Their Applications*. Algèbre non commutative, groupes quantiques et invariants (Reims, 1995), 263–279, Sémin. Congr., 2, Soc. Math. France, Paris, 1997. [MR1601155].
- [N] A. Nowicki, *Integral Derivations*. J. Alg. **110** (1987), no. 1, 262–276.
- [Z1] W. Zhao, *Images of Commuting Differential Operators of Order One with Constant Leading Coefficients*. J. Alg. **324** (2010), no. 2, 231–247. [MR2651354]. See also arXiv:0902.0210 [math.CV].
- [Z2] W. Zhao, *Generalizations of the Image Conjecture and the Mathieu Conjecture*. J. Pure Appl. Alg. **214** (2010), 1200–1216. See also arXiv:0902.0212 [math.CV].
- [Z3] W. Zhao, *Mathieu Subspaces of Associative Algebras*. J. Alg. **350** (2012), no.2, 245–272. See also arXiv:1005.4260 [math.RA].
- [Z4] W. Zhao, *Some Open Problems on Locally Finite or Locally Nilpotent Derivations and \mathcal{E} -Derivations*. Preprint.
- [Z5] W. Zhao, *Idempotents in Intersection of the Kernel and the Image of Locally Finite Derivations and \mathcal{E} -derivations*. Preprint.
- [Z6] W. Zhao, *The LFED and LNED Conjectures for Algebraic Algebras*. Preprint.
- [Z7] W. Zhao, *Images of Ideals under Derivations and \mathcal{E} -Derivations of Univariate Polynomial Algebras over a Field of Characteristic Zero*. Preprint.

DEPARTMENT OF MATHEMATICS, ILLINOIS STATE UNIVERSITY, NORMAL, IL 61761. EMAIL: WZHAO@ILSTU.EDU